基于关联数据的历史档案资源聚合研究

■ 贾琼　王萍
吉林大学管理学院　长春 130022

摘　要：[目的/意义] 应用关联数据技术进行历史档案资源聚合，以发掘资源数据的潜在关联，发现新知识并实现价值增值，进而优化知识服务。[方法/过程] 梳理档案资源开发利用和聚合的相关研究，从资源聚合的目标、原则、流程和层级关系等方面构建历史档案资源聚合框架，并采用"经史子集"实例分析资源聚合过程。[结果/结论] 构建了关联数据视角下历史档案资源聚合的理论指导框架和检索服务平台，为历史档案资源的开发利用提供从理论到实践的实现路径。

关键词：关联数据　历史档案　资源聚合　本体　元数据
分类号：G250
DOI：10.13266/j.issn.0252－3116.2021.10.011

1　引言

大数据时代，网络信息资源规模海量增长，图书馆、博物馆、档案馆等文化机构储藏了大量的历史档案资源，而历史档案资源的数字化保护与开发利用还需不断地纵深发展，以满足用户深度获取资源的需求。近年来，关联数据的广泛应用与发展使各信息资源部门逐步实现了 数据网络的构建。关联数据采用三元组方式对不同形态数字资源进行解析，并协同本体进行深层次的描述，最大限度地揭示了数字资源的内部关联，进而实现数字资源间的互操作。关联数据具有针对解决当前历史档案资源深度组织与开发中所存在的信息孤岛和信息超载难题的潜力。资源聚合中的"推理功能"更有助于发现潜在的新知识链接。本研究提出基于关联数据的历史档案资源聚合思路，以便加深对历史档案资源聚合概念的理解，推动聚合技术的深入研究，加快实现历史档案资源的互联与聚合进程。

2　相关文献研究

2.1　历史档案资源开发利用

对于历史档案资源开发利用，学者们在理论方法及策略构建方面积累了大量的有益成果，同时为技术应用于历史档案资源开发利用的微观实践打下了坚实基础。L. Y. Chen 等[1]以非遗为研究对象，主张建立与非遗特征相契合的管理信息系统，在跨学科理念视角下搭建领域互通的桥梁。牛力等[2]从数字人文研究一般方法论体系和数字环境下档案资源开发利用方法体系中抽离组配逻辑，结合档案资源的基本特征，形成了包含发现、重构和故事化的方法路径体系。赵雪芹等[3]从转变数字档案信息资源整合方式、扩大档案信息资源开发工作的规模、深化档案部门与数字人文研究团队的合作、拓宽档案信息资源开发涉及的领域 4 个方面构建了基于数字人文的档案信息资源开发的路径体系。陈佳雨等[4]构建了档案故事众开发模式，并从资源层、转化层与传播层提出开发策略。朱兰兰等[5]从基于文本挖掘技术的家谱档案资源深度开发、基于数据管理技术的新型家谱档案数据库建设、基于 VR/AR 技术的家族档案和场景重构、基于 GIS 技术的家谱资源和文化共享平台建设 4 个数字人文的理论方法维度思考了家谱档案资源多元化开发利用的可行思路。学者们还分别以地方特色档案[6]、黄河档案资源[7]、口述历史档案[8]等为研究载体，利用数字人文方法，提出了多元化的历史档案资源开发方式。

2.2　资源聚合

资源聚合研究在图书情报学领域研究较为丰富。T. Heath 等[9]通过构建 Web 数据和关联数据应用程序，来加快全球范围内数据集成的实现。邱均平等[10]

作者简介：贾琼（ORCID：0000－0002－2696－7282），博士研究生；王萍（ORCID：0000－0002－0308－7362），教授，博士生导师，通讯作者，E-mail：jlwang65@163.com。
收稿日期：2021－03－02　修回日期：2021－04－09　本文起止页码：105－112　本文责任编辑：易飞

以情报领域为研究范围研究作者共被引的情况,研究表明基于共被引的馆藏资源聚合是一种良好方法。王洪军等[11]在对元数据进行采集、整合与存储后,构建了基于元数据仓储和动态链接的资源整合系统,以便用户通过知识导航、信息检索、个性化服务等功能实现对馆藏资源的检索与使用。游毅等[12]用关联数据实现馆藏资源的语义关联、资源浏览等服务,并在图书馆关联数据链接管理中通过建立 RDF 链接创建与维护机制实现与馆外数据的关联。瞿辉等[13]对语义共词分析模型进行构建,根据结果对馆藏资源进行语义标注,从而实现馆藏资源的多维聚合。从一定程度上来说,档案资源聚合也已进入档案学界的研究视野,但研究步伐却相对缓慢。张云中等[14]以 NARA 数字档案标注系统中的"Women at War"为案例,构建了数字档案资源聚合模型,有助于提升数字档案资源的服务效益。舒忠梅[15]提出了档案时空本体模型、档案数据抽取框架并构建出档案关联数据知识图谱,以拓展开发和创新利用档案馆藏资源,实现档案资源增值。

2.3 研究述评

通过系统梳理国内外研究现状,发现学者们在历史档案资源开发利用的方法路径及策略构建方面形成了较为系统研究体系,对该领域的发展起到了极大的推进作用。但是,总的来说当前研究体系中宏观理论较多,技术应用较少。国内外已在多重领域进行了基于关联数据的数字资源聚合研究,尝试将关联数据与非物质文化遗产、图书情报学、政府等互联,实现资源深度聚合。而档案学界将关联数据技术应用到历史档案资源聚合研究的步伐要相对缓慢。

在信息技术日益更迭的数字时代,历史档案资源的开发也迎来了资源分散、异构、信息孤岛等一系列理论及应用难题。而关联数据则为历史档案资源的聚合提供了一个可行的研究思路和方案,不仅有助于解决海量分散资源的异构甚至信息孤岛等问题,还在推动历史档案资源的开发进程、丰富历史档案资源理论、拓宽历史档案资源聚合的技术路线等方面发挥着极其重要的作用。基于此,本研究以海量的、分散的、异构的、价值丰富的历史档案资源为研究对象,引入关联数据的方法、技术与理念,以实现对历史档案资源的深度聚合。

3 历史档案资源聚合分析框架

构建基于关联数据的聚合框架模型是实现历史档案资源聚合的重要基础。以下通过探索历史档案资源聚合的目标、原则和流程,为构建基于关联数据的历史档案资源聚合框架模型提供依据。

3.1 资源聚合目标

历史档案资源聚合的总体目标是着眼于用户所需的资源集,以知识体系的方式呈现给用户,改变传统的独立知识片段的呈现结果,节约用户的检索时间,尽可能多地满足用户的知识需求,最大化挖掘历史档案资源的利用价值,提升知识服务的效能[16]。具体包括以下目标:

3.1.1 资源组织

历史档案资源的载体有纸质、照片、录音、录像、电子等,种类繁多,且资源结构和类型各不相同。数据异构导致历史档案资源的利用、开发和共享的效果并不理想。将历史档案资源进行资源的聚合研究,终极目标便是将资源数据库中的异构数据信息进行集中聚合,从而推进信息的流通和共享,这里的资源聚合不仅仅是将分散资源简单地整合到一起,而是以整合为基础,先合理规范资源,再将其进一步划分成为不同的模块[17]。历史档案资源聚合的最终结果是要打破原有的资源分散、独立、单一的模式,重塑资源的结构体系,构建全新的、规范化的、成体系的资源结构体。

3.1.2 知识服务

在历史发展与研究中,如何有效聚合历史档案资源是本文需要解决的关键问题。在当前环境中,专家学者们运用了大量的本体构建、关联数据、语义聚合等手段对历史档案资源进行分析,期望能够有效分析同一历史档案资源中不同主题的分布态势。历史档案资源及其实体的聚合能为历史主题和研究结构的深度探索提供资料和素材,加上技术的辅助支撑,比如在历史主题和研究结构的研究中采用科学地图和可视化等技术的辅助支持,借助可视化的呈现结果,结合聚类等方式,专家学者们可以快速理清历史研究的结构[18]。

3.1.3 科学交流

科学交流在学术界是一种较为大型、普遍的社会现象,其最为典型的表现是科学合作。传统的以个人搜集等方式获取的历史档案资源存在各种难以避免的问题。相比传统的历史档案资源数据获取方式和研究方法,历史档案资源之间的学术关联网络,为资源网络研究提供了良好的研究平台,有助于科研工作者更好地进行学术交流。

3.1.4 学科交叉

在历史研究主题结构中,采用跨学科、跨领域的方式进行研究,是该研究中一直被重点关注的热门话题。

贾琼，王萍. 基于关联数据的历史档案资源聚合研究[J]. 图书情报工作,2021,65(10):105－112.

通常来说，从事历史档案研究的专家学者们按照一定的资源分类依据，圈定某个历史领域的资源，进而针对资源包含的实体对象如历史人物或者史料，借助定量分析方法分析不同类型的实体之间的关联或者相互被引情况，借此可看出跨学科之间的分析交流状况。

3.2　资源聚合原则

本文参考档案元数据和 DC 元数据等相关的标准与规范，建立历史档案资源元数据标准规范，以满足资源实体和数字对象的定义、修复、组织、解译、管理、保管和维护等的需要，维持历史档案资源的生命周期活动。聚合过程中应注意以下几点：①需要聚合的历史档案资源应为异构的、单独存在的和分布式的；②资源聚合的第一步是聚集，第二步才是聚合，即先将现有的历史档案资源组织成一个整体，而后再对资源进行序化，实现聚合；③聚合过后的系统要能够为历史档案研究者和相关学者们提供一个可供检索的平台；④聚合的最终目的是提高资源利用率，便于历史档案相关研究的开展和深入。因此，历史档案资源的聚合不是简单地把各种不同类型和来源的异构历史档案资源聚合到一个相同的 Web 网络环境中，更重要的是要实现历史档案资源服务与资源的深度聚合。

3.3　资源聚合流程

通过聚合，可将历史档案资源划分为各类实体，分析其详细的属性与关系，组织资源目录，赋予目录关联与检索功能，比如层次清晰地展示与某朝代、历史人物或文献相关的所有资源，关联某个历史档案资源所涵盖的所有内容表达、载体表现。基于关联数据的资源语义聚合，能够揭示历史档案资源表层内容的结构及深层次的知识单元，并重点关注资源实体之间的相互联系，建立知识网络。

历史档案资源聚合过程涵盖了采集、资源描述、语义聚合及可视化呈现等阶段，见图 1。

3.3.1　资源采集阶段

历史档案资源数据包含不同类型且来源于不同系统，在存储分布上会出现异构异质的特点。在有针对性地收集历史档案资源时，应该明确历史档案资源的特征和来源可靠性，历史档案资源必须是真实有效的环境所产生的。

3.3.2　资源描述阶段

资源描述是资源聚集的前提和基础，需要遵守统一的元数据标准和框架。为方便计算机识别和处理，本文复用较为成熟的元数据标准体系统一描述不同来源、不同结构、不同类型的数据资源。

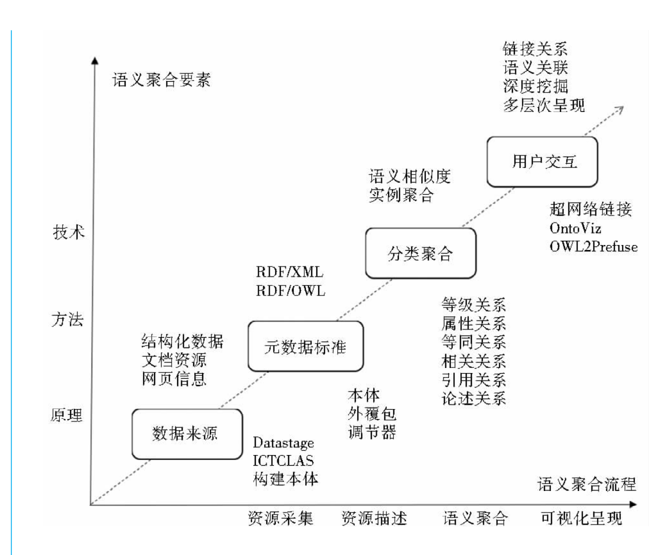


图 1　基于关联数据的历史档案资源聚合流程

3.3.3　语义聚合阶段

整个流程的核心是语义聚合。聚合历史档案资源语义时可以：①利用先验语义进行语义分类聚合，如根据主题或内容的语义相似度实现相同语义特征的资源的聚集；②根据本体概念及概念关系的组织进行资源聚合，如揭示资源实例元素的人名、机构名、题名等元数据间的语义关系并进行语义关联及映射，从而实现资源聚合。

3.3.4　可视化阶段

以视觉的方式呈现出来的结果就是可视化，一方面可以直观地展现历史档案资源中复杂的链接和关联；另一方面，在大量的数据与复杂的关联中，用户可以利用直观的认知和理解能力发现其中的潜在知识。根据历史档案资源的重要程度来设定相应的表现形式和视觉距离，要符合大众用户的视觉思维和习惯，是可视化需要的结果。历史档案资源聚合后的可视化呈现，有利于用户的感觉、知觉共同调动，从而协助用户发现知识。

4　基于关联数据的历史档案资源聚合的实现

为使历史档案资源具备语义性，能够与其他资源广泛链接，在本文中笔者利用关联数据技术对其进行重组。因此，历史档案资源内容建设是本节研究的重点，下文从历史档案资源聚合层级关系及框架提出基于关联数据的历史档案资源聚合方法。

4.1　历史档案资源聚合层级划分

本文从本体、内部、外部的资源聚合 3 个层级探索历史档案资源聚合，如图 2 所示：

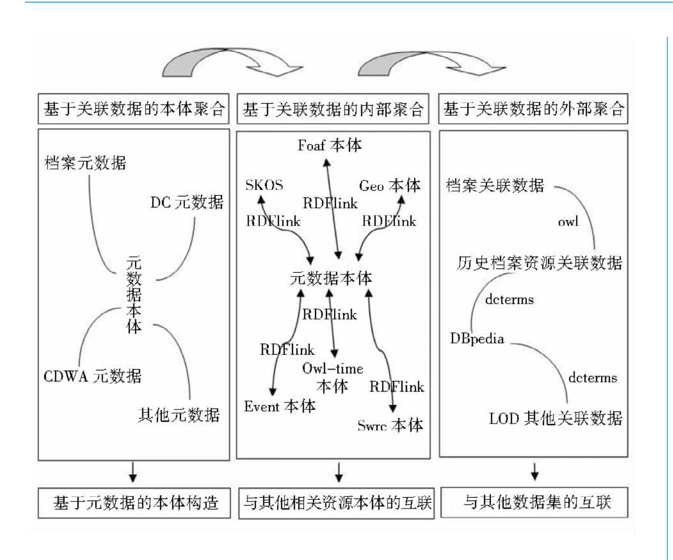


图 2 历史档案资源聚合的层级关系

4.1.1 基于本体的资源聚合

历史档案资源由不同来源、不同类型的档案资源组成，为了便于统一描述资源，需要制定统一的元数据标准格式，以实现元数据之间的互操作。参照 OAI 元数据收割协议[19]等，将不同来源、不同类型、不同格式的异构元数据(MARC、EAD、LIDO、TEI 等)与元数据本体间建立映射，获得符合描述规范的数据集，利用 XSTL 将数据集进行 XML 转换，生成符合 RDF/XML 描述规范的元数据，实现异构元数据的语义聚合。

4.1.2 基于关联数据的内部聚合

根据关联数据发布原则，利用 URI 命名历史档案资源，并描述资源语义。借助 OCLC 提供的语义相似性算法，计算以 RDF 格式表示的历史档案资源之间的相似度，挖掘具有各种隐性关系的相关资源，并综合利用 SKOS、VIAF、GEO、FRBR、FOAF[20]等被广泛应用的词库映射，将这种隐性的语义关系转化为显性的 RDF 语义关系。RDF 可以链接异构资源形成知识网络，以此完成领域内部资源聚合。

4.1.3 基于关联数据的外部聚合

历史档案资源发布需要内部与外部数据集关联，内外部数据集来源于多个数据渠道。通过 URI 来识别和定位 RDF 描述资源和相关关联资源的 RDF 链接，并使用语义标注工具 Annocultor tool[21]对 URI 标识的关联数据进行标注，得到关联数据的 Web 参引解析 HTTP URI，并以 HTTP URIs 为桥梁，通过 LOD 云(Liked Open Data Cloud，关联开放数据云)提供的服务与外部相关的开放关联数据集语义互联的服务，实现历史档案资源与外界资源的跨域聚合。

4.2 历史档案资源聚合结构框架

基于关联数据的历史档案资源聚合框架分为：数据层、发布聚合层、服务应用层，如图 3 所示：

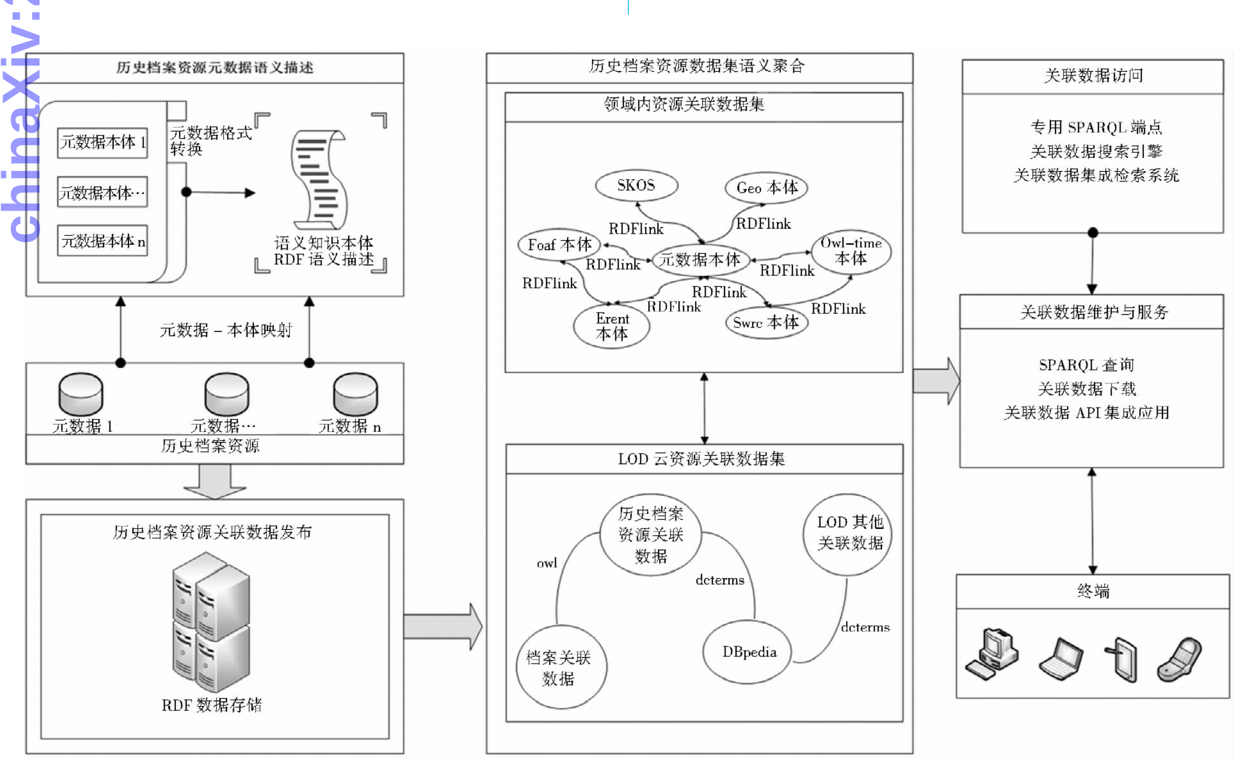


图 3 基于关联数据的历史档案资源聚合框架

4.2.1　语义模型构建

对于历史档案资源的内容描述,在不同的历史阶段有不同的元数据标准,为让不统一的历史档案资源描述信息能够拥有一致的管理标准和模式,参考 METS 和 RDF 等有关规定,构建了开放的模型,此模型聚合了多种元数据标准[22]。

主要包含的资源核心类有:描述不同数据源的数据集合、数据集合中包含的历史档案资源对象、资源对象的数字表现形式。为使不同类型、来源和格式的历史档案资源的语义描述相一致,依据 ORE 的标准,采用 ore:Aggregation 类关联全部的历史档案资源对象和语义描述,使其形成一个数据集合网络。另外,为强化描述中的关联,此模型还创建了查找历史档案相关资源的元数据属性,包含名称、标识、版本、定义等,通过使用统一的描述标准,关联资源对象,使其更好地聚合。

4.2.2　关联数据的发布与聚合

模型建立后,首先是本地资源的数据关联,其次通过词表映射、实体识别、数据聚合等操作,建立历史档案资源语义服务的关联数据集,最后定义词汇术语和本体描述数据信息的概念、术语和逻辑结构。本文通过建立数据集之间关联关系实现的聚合,可以实现数据集之间的资源共享,该过程可以描述为:①通过词表映射找到同一概念或关系的不同词汇描述;②通过实体识别在不同的数据集中建立同一实体数据之间的关联;③利用聚合过程,通过人工干预对同一实体对象的冲突属性进行验证、过滤和合并。

4.2.3　聚合服务

在这个层级上,不仅可以对历史档案资源数据进行语义链接,还可以为用户提供资源访问和检索服务。历史档案资源聚合服务的核心价值是提高数据质量,提供具有独特见解的结构化经验和信息。利用具有语义支持的 Web Service 模型对历史档案资源进行知识表示、描述和重组,实现历史档案资源数字数据的再利用和共享,提高其使用价值。

5　历史档案资源聚合的应用实例分析

为了提高历史档案资源聚合框架的可靠性和可用性,在此采用“经史子集”实例 验证框架,通过利用 protégé 和 D2R 分别对历史档案资源进行本体构建和数据集发布。

5.1　元数据方案

历史档案资源的种类复杂,数字资源的描述不但要揭示对象本身还要对历史档案进行互联,因此,相对于其他单个信息资源而言,描述具有一定的难度。本文中历史档案资源元数据核心元素参考了 FRBR、CDLS、DC、CDWA 等标准,结合历史档案资源与档案数据的属性,加以横向延伸及纵向拓展,共分为 4 类、21 个元素,如表 1 所示:

表 1　历史档案资源核心元数据

分类	元素
历史档案描述信息	名称、年代、类别、标识符、主题、来源、创作、规格、位置、描述
数字对象管理信息	创建者、时间、版权
数字对象描述信息	文件名、大小、格式、类型
相关资源	相关人物、相关知识、相关事件、关联数字资源

5.2　本体构建

本文依据七步法[23]和 Protégé 技术构建经史子集领域本体,具体实现过程如图 4 所示:

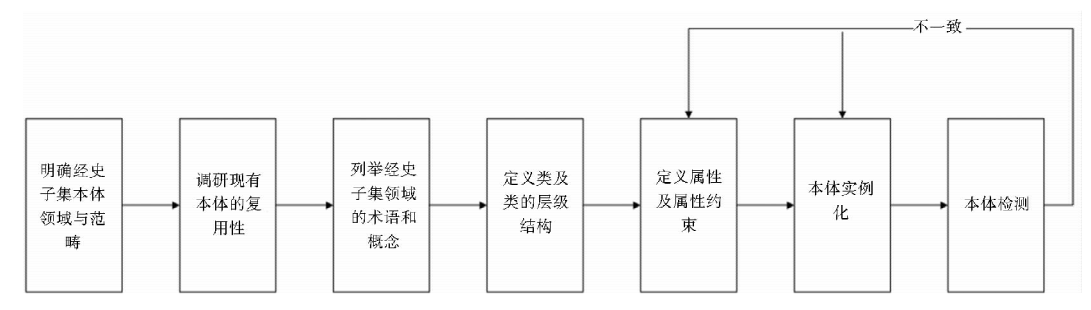


图 4　经史子集领域本体构建流程

(1)明确本体的领域和范畴,构建的本体是针对历史档案资源中的经史子集类别。

(2)复用已有本体,以提高工作效率,获得更多交互。通过对文献进行查阅调研之后,发现已有的经史子集元数据大多是依据 2006 年修订的《古籍著录规则》进行著录的,统一编目、分类,对主要信息进行充分揭示[24]。

(3)列举经史子集领域的术语和概念,尽可能多地获取相关概念和术语,为后面定义层级关系奠定基础。

(4)定义类及类的层级结构,从经史子集领域术

语中选取相关词汇,并研判词汇间的逻辑关系,以此构建层级结构。

(5)定义属性及属性约束为层级关系的概念框架,因此若要明确属性与类间的关系,需要添加属性描述。在 Protégé 中可以为 Class、String、Number 等类型。经史子集部分数据与对象属性见图 5。

为了让本体语义关系能够被更清晰、完善地表达,定义了概念语义关系后,对本体进行可视化,如图 6 所示。本体可视化与经史子集类层次相比能体现更加丰富具体的联系描述。



图 5 经史子集本体数据与对象属性(部分)

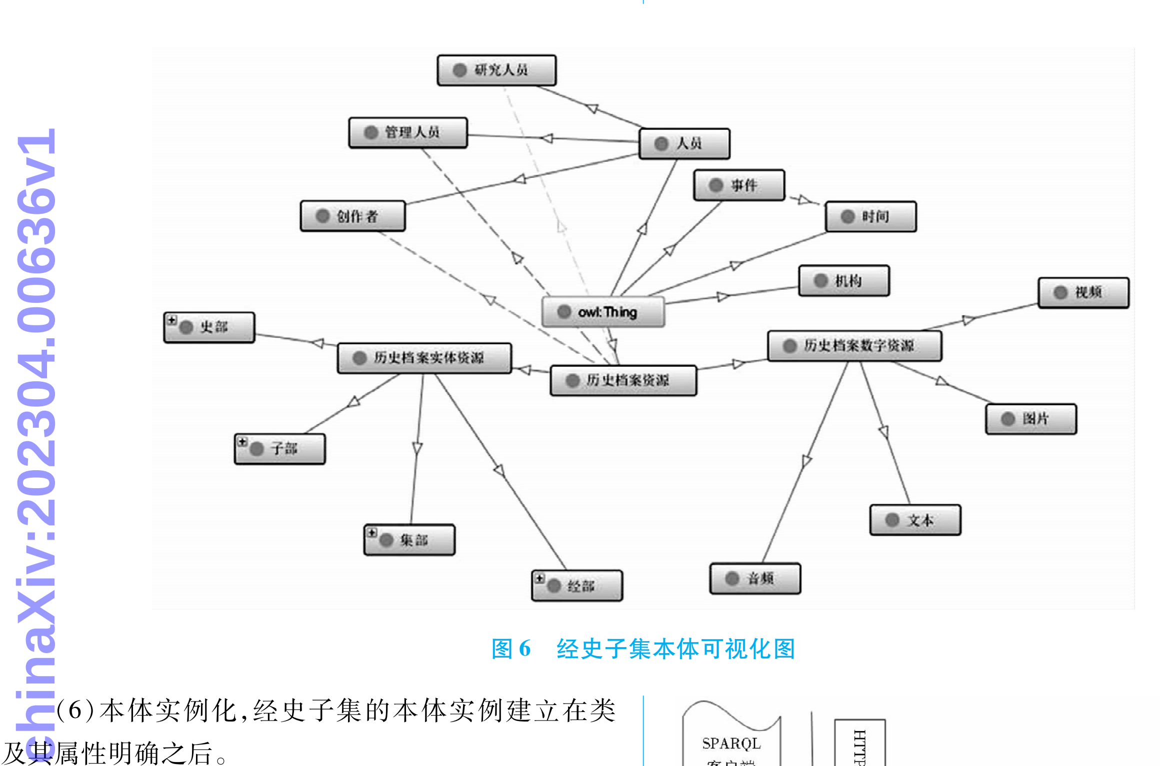


图 6 经史子集本体可视化图

(6)本体实例化,经史子集的本体实例建立在类及其属性明确之后。

(7)本体检测,验证本体是否符合逻辑,包括类、属性、实例的合理性问题。

5.3 关联数据发布

现今的门户网站与信息系统大多以关系型数据库对数据进行存储,因此应将关系型数据库转化为 RDF 形式。随着关联数据的发展与应用,已有学者研发出如 D2R、Squirre1RDF 及 Virtuoso 等将数据库中的数据转化为关联数据集的工具,以 D2R 为例,其体系结构见图 7。

D2R 主要提供 RDF 数据的查询接口,RDF 存储、HTML 浏览器和 SPARQL 客户端均可调用。D2R 主要包括 D2R 服务器、d2rq 引擎和 d2rq 映射语言。其主要功能是使用可定制的 d2rqmapping 将数据库中的数据转换为虚拟 RDF 格式的数据,并将数据转换为真实的 RDF 文件。

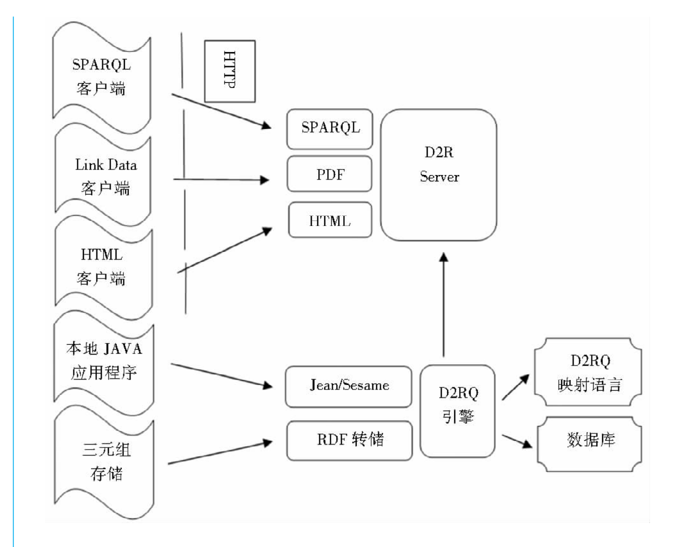


图 7 D2R 的结构体系

D2R 自带的启动服务程序能够为我们提供使用,启动 d2r. server. bat,只需要输入命令"generate-mapping -u root -p 123123 -o db_relic. ttl jdbc:mysql:///db_rel-

ic? useSSL = false"执行已经生成的 mapping 文件,然后运行 D2R,显示[[[Server started at http://localhost:2020/]]]就可以在浏览器中进入到 D2R 服务页面。

5.4 资源聚合

由于 D2R 发布关联数据集的检索方式只能通过 SPARQL 的方式去检索,普通的检索用户并不适用,本文采用 MYSQL + ECLIPSE 的开发技术,设计了历史档案资源经史子集数字服务平台来对资源进行聚合。

该平台为用户提供主题、年代、地点等检索方式,如在检索"左传"时,可以按照"左传""左丘明""春秋左氏传"等相关关键词进行检索,在检索结果中,会显示名称、图片、主题、人物等相关信息。在双击名称后,平台为用户提供了相关研究学者的研究成果,主要是研究论文。由于海量的历史档案资源可能存在内在联系,可以通过相关事件及人物进行点击查询,显示相关事件、人物的检索资源,如图 8 所示:



图 8 经史子集数字资源服务平台检索结果页面

6 结语

数字人文视域下,发展、挖掘和应用传统文化资源成为了研究热点。为了发现历史档案数字资源的隐性数据关系、潜在知识和增值价值,需借助人文计算工具。关联数据对资源实体标识了唯一的 URI,并利用 RDF 链接数字资源,从而形成数据网络。本文从元数据描述、元数据 – 本体映射、RDF 数据发布、语义聚合与访问的层面构建了历史档案资源聚合框架,该框架为历史档案数据资源聚合提供了理论引领。历史档案资源聚合框架给"经史子集"实例的资源聚合提供了实现的路径体系,最终建立的历史档案数据资源聚合服务平台可满足用户查询和检索历史档案的需求。

参考文献:

[1] CHEN L Y, WANG B Z, CHEN B. Research on digital museum for the intangible cultural heritage[C]//IEEE international symposium on IT medicine and education. Jinan,2009:734 – 737.

[2] 牛力,高晨翔,张宇锋,等. 发现、重构与故事化:数字人文视角下档案研究的路径与方法[J]. 中国图书馆学报,2021,47(1):88 – 107.

[3] 赵雪芹,党昭,李天娥. 数字人文视角下的档案信息资源开发问题与对策[J]. 北京档案,2021(1):18 – 22.

[4] 陈佳雨,汤玲玲,王小云. 自媒体时代档案故事众开发模式研究[J]. 浙江档案,2021(1):28 – 30.

[5] 朱兰兰,薄田雅. 数字人文视域下家谱档案资源多元化开发[J]. 浙江档案,2019(10):31 – 33.

[6] 卢丹丹,聂云霞. 数字人文视角下地方特色档案资源开发路径[J]. 山西档案,2020(4):29 – 33.

[7] 李宗富,于子闪. 数字人文视域下的黄河档案资源开发利用研究[J]. 山西档案,2021(1):73 – 79.

[8] 高淞,王向女. 数字人文视域下口述历史档案资源开发利用研究[J/OL]. 山西档案. [2021 – 04 – 09]. https://kns. cnki. net/kcms/detail/14. 1162. G2. 20210112. 1728. 002. html.

[9] HEATH T, BIZER C. Linked data:evolving the Web into a global data space[J]. Synthesis lectures on the semantic Web: theory and technology,2011,1(1):1 – 136.

[10] 邱均平,周毅. 基于作者共被引的馆藏资源深度聚合模式与服务探析:以 CSSCI 中图书情报领域本体研究为例[J]. 图书情报工作,2014(7):19 – 24.

[11] 王洪军,刘明,李焱,等. 基于元数据仓储与动态链接的图书馆资源整合系统的整合与实现[J]. 中华医学图书情报杂志,2011(10):65 – 67,70.

[12] 游毅,成全. 试论基于关联数据的馆藏资源聚合模式[J]. 情报理论与实践,2013(1):109 – 114.

[13] 瞿辉,邱均平. 基于语义化共词分析的馆藏资源聚合研究[J]. 情报科学,2016,34(2):15 – 20.

[14] 张云中,冯双双. 基于主题图的数字档案标注系统资源聚合研究[J]. 图书情报工作,2018,62(14):116 – 124.

[15] 舒忠梅. 数字人文背景下的档案知识图谱构建研究[J]. 山西档案,2020(2):53 – 60.

[16] 董克. 数字文献资源多元深度聚合研究[D]. 武汉:武汉大学,2014.
[17] 杨现民. 泛在学习资源动态语义聚合研究[J]. 电化教育研究,2014,35(2):68-73.
[18] 张萌. 基于少数民族档案关键词的社会网络分析[J]. 黑龙江档案,2019(2):23-24.
[19] 丁遒劲,苏静,曾建勋. 国家元数据库及其协同构建框架研究[J]. 情报理论与实践,2020,43(10):88-92,80.
[20] 张鹏图. 大英图书馆书目数据的关联化分析[J]. 国家图书馆学刊,2015,24(4):103-113.
[21] 王萍,黄新平. 基于关联开放数据的数字文化资源语义融合方法研究——欧洲数字图书馆案例分析[J]. 图书情报工作,2016,60(12):29-37.
[22] Europeana data model primer[EB/OL]. [2021-04-09]. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf.
[23] 李大鹏. 基于本体的学科知识地图构建研究[D]. 武汉:华中师范大学,2011.
[24] 屈亚文. 基于关联数据的古籍数字资源组织[J]. 百科知识,2020(27):5-7.

作者贡献说明:
贾琼:负责论文撰写与修改;
王萍:负责研究思路框架设计。

Research on Aggregation of Historical Archive Resources Based on Linked Data

Jia Qiong Wang Ping
School of Management, Jilin University, Changchun 130022

Abstract: [**Purpose/significance**] The application of linking data technology in the aggregation of historical archive resources can help to explore the potential correlation of resource data, discover new knowledge and realize value increment, so as to optimize knowledge service. [**Method/process**] This paper sorted out the related researches on the exploitation and utilization of archival resources, constructed the framework of archival resources aggregation from the objective, principle, process and hierarchical relationship of resource aggregation, and analyzed the process of resource aggregation with the example of "Jing Shi Zi Ji". [**Result/conclusion**] The theoretical guiding framework and retrieval service platform of historical archival resources aggregation from the perspective of linked data are given, which provide a path from theory to practice for the development and utilization of historical archival resources.

Keywords: linked data historical archive resource aggregation ontology metadata